ED276097

RESEARCH roundup

Volume 2, Number 2

April 1986

ED276097

U.S. DEPARTMENT OF EDUCATION

ORIGINAL

EM 018 234

DOCUMENT RESUME

ED 276 097 EA 018 284

AUTHOR	Ellis, Thomas I.
TITLE	Teacher Evaluation.
INSTITUTION	ERIC Clearinghouse on Educational Management, Eugene, Oreg.; National Association of Elementary School Principals, Alexandria, VA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Apr 86
CONTRACT	400-83-0013
NOTE	6p.
AVAILABLE FROM	Publication Sales, National Association of Elementary School Principals, 1615 Duke Street, Alexandria, VA 22314 ($2.00; quantity discounts; Virginia residents should add 4 percent sales tax).
PUB TYPE	Information Analyses - ERIC Information Analysis Products (071) -- Collected Works - Serials (022) -- Reference Materials - Bibliographies (131)
JOURNAL CIT	Research Roundup; v2 n2 Apr 1986
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Administrator Role; Concurrent Validity; Cooperative Planning; *Decision Making; Educational Environment; Elementary Secondary Education; Employment Practices; *Evaluation Methods; Instructional Improvement; Legal Responsibility; Organizational Climate; Performance Factors; *Personnel Policy; Program Development; *Program Effectiveness; Program Implementation; *Teacher Administrator Relationship; Teacher Attitudes; Teacher Effectiveness; *Teacher Evaluation; Teacher Improvement; Teacher Supervision

ABSTRACT

Included in this summary analysis of three journal articles and two documents on teacher evaluation are reports that touch on difficulties facing school officials in developing a teacher evaluation system and suggestions for assessing the strengths and weaknesses of an evaluation system. The first article, by Susan S. Stodolsky, challenges evaluation methods that rely on classroom observation alone and emphasizes the importance on the entire instructional context. The second, a document written by Rand Corporation researchers, concludes from a study of four exemplary school districts that the teacher evaluation process is inseparable from the school district's larger organizational context. The three remaining selections focus on practical aspects of implementing a teacher evaluation system. In his guide, Thomas L. McGreal identifies nine "commonalities" of effective systems that support his conviction that evaluation should help, rather than judge, teachers. Taking the position that some form of summative evaluation is nonetheless needed for making personnel decisions, James Raths and Hallie Preskill offer recommendations to guide administrators through the task of teaching staff assessment. Because of the potential legal implications brought with any personnel decision, the final selection by Donovan Peterson provides guidelines for ensuring that a teacher evaluation system will withstand judicial scrutiny. (CJH)

roundup

Volume 2, Number 2 April 1986

Teacher Evaluation

Thomas I. Ellis

The current educational reform movement has thrown a spotlight on teacher competency, leading policy makers to pursue a variety of measures ranging from merit pay and career ladders to tenure review and mandatory dismissal of those teachers who are demonstrably incompetent. All such proposals inevitably underscore the need for an effective, reliable, and legally defensible system of teacher evaluation.

Teacher evaluation need not be what it too often becomes: an essentially meaningless formality regarded with suspicion and even contempt by teachers and as frustrating by supervisors. Moreover, it does not have to be a source of contention between teachers and administrators. If a teacher evaluation system is research-based, designed to improve instruction, and approached with a cooperative attitude by all parties, it can be an effective and dynamic agent for educational renewal.

In developing a teacher evaluation system, school officials must confront two fundamental difficulties. The first is that no completely objective approach to assessing teacher performance has ever been found (though in recent years researchers have been making gains on this problem). Indeed, it has become axiomatic that the more explicit or standardized the criteria become, the less valid they are for assessing a particular teacher's effectiveness.

The other problem is that the two major purposes of teacher evaluation—helping teachers to become better at their job and providing a basis for making personnel decisions—are usually perceived as being at odds with each other. Effective formative evaluation requires a relationship marked by mutual trust between teachers and supervisors. Yet such a relationship is hardly possible if the teacher feels that the evaluation process may lead to an adverse personnel decision or if the supervisor fears that such a decision may lead to a court suit.

Five recent reports that touch on these difficulties provide useful guidelines for assessing the strengths and weaknesses of just about any school's teacher evaluation system. The first, by Susan S. Stodolsky, challenges the validity of evaluation methods that rely on classroom observation alone and emphasizes the importance of looking at the whole context of instruction (subject matter and activity structure as they relate to teaching style) when evaluating teacher performance.

In the second, several Rand Corporation researchers concluded from their study of four exemplary school districts that the teacher evaluation process is inseparable from the larger organizational context of the school district. Findings from the study suggest that organizational commitment and staff involvement are essential for effective teacher evaluation, regardless of the methods used.

The remaining three selections focus on practical aspects of implementing a teacher evaluation system. In his comprehensive guide based on extensive experience, Thomas L. McGreal identifies nine "commonalities" of effective systems, all in keeping with his conviction that the primary purpose of evaluation should be to help, rather than to judge, teachers.

Turning to the proposition that some form of summative evaluation is nonetheless needed for making personnel decisions, James Raths and Hallie Preskill offer some recommendations to guide administrators through the delicate task of deciding which members of the teaching staff are excellent, satisfactory, or incompetent.

Because any personnel decision based on teacher evaluation is fraught with potential legal troubles, the final selection by Peterson offers practical guidelines for ensuring that a teacher evaluation system will withstand judicial scrutiny.

Thomas I. Ellis, Ph.D., is a research analyst and writer at the ERIC Clearinghouse on Educational Management, University of Oregon.

Prepared by ERIC Clearinghouse on Educational Management

NAESP

NATIONAL ASSOCIATION OF ELEMENTARY SCHOOL PRINCIPALS

ED 018 284

1 Stodolsky, Susan S. "Teacher Evaluation: The Limits of Looking." *Educational Researcher* 13, 9 (November 1984), 11-18. EJ 309 391.

Most teacher evaluations rest on a few direct observations, usually unspecified as to situation. The basic assumption is that effective teaching can be reduced to a set of generalized behaviors that are consistent across teaching situations and occasions.

Stodolsky challenges this assumption. Wielding both conceptual arguments and empirical data, she shows that the effectiveness of any teaching style or behavior depends on the context in which the teaching takes place. Hence, evaluations based on a small number of classroom observations cannot do justice to the range of teaching behaviors, skills, and arrangements used for different purposes by teachers, particularly at the elementary level.

Stodolsky proposes a flexible "activity structure" approach, tied to subject matter and curriculum, as a more accurate characterization of elementary school teaching. The evaluator assesses how appropriate the teaching style and approach are to the subject matter at hand and to the skill level of the students. To do this, the evaluator must describe how activities are structured, who is present, how long the activity lasts, and what its instructional purpose or format is.

The unit of analysis in this method is not the usual random selection of different classes, but rather what Stodolsky labels "activity segments." Each activity segment comprises a particular set of instructional formats, participants, materials, behavioral expectations and goals, and time boundaries. The evaluator takes all of these variables into account in determining the appropriateness or effectiveness of any given instructional strategy. Such an approach does justice, it is pointed out, to the nature of teaching as a context-bound activity.

A study of fifth-grade math and social studies classes in Chicago tested Stodolsky's proposed approach. The findings indicate that one should expect systematic variation in teaching and instructional arrangements, not consistency. The same teachers teaching the same children in the same physical setting used very different instructional arrangements as they switched from math to social studies. Mathematics instruction proved to be homogeneous within and across classrooms, whereas social studies instruction characteristically tends to be diverse. Stodolsky proceeds to describe in detail three key features of the activity segments studied: instructional format, pacing, and cognitive level.

The data indicate that elementary school teachers are essentially generalists who create a broad repertoire of organizational and pedagogical arrangements. Thus assumptions of internal consistency made in connection with teacher evaluation procedures must seriously be questioned. Stodolsky therefore concludes that evaluators should examine teaching within an overall context—taking into account the subject matter and the structure and purpose of classroom activities—rather than simply identifying the presence or absence of a list of teaching "behaviors" judged to be desirable. She concedes, however, that such a complex and extensive task for the study of teachers and teaching will not win easy converts.

2 Wise, Arthur E., and others. *Teacher Evaluation: A Study of Effective Practices.* Santa Monica, California: Rand Corporation. Sponsored by the National Institute of Education, Washington, DC, June 1984. 101 pages. ED 246 559.

Teacher evaluation does not occur in a vacuum. Wise and his colleagues note: it is shaped by the organizational, political, and instructional context in which it takes place. So how do these contextual factors influence the quality of teacher evaluation programs? To find out, the Rand Corporation conducted case studies in four school districts with highly effective, but very different, teacher evaluation programs: Salt Lake City, Utah; Lake Washington, Washington; Greenwich, Connecticut; and Toledo, Ohio. Despite their varying approaches, certain common characteristics set these systems apart from less successful ones.

The first common factor was organizational commitment—the willingness of the districts' top-level leadership to devote adequate time, personnel, and institutional resources to teacher evaluation. The second common factor was the recognized competence of the evaluators in making judgments and recommendations, coupled with mechanisms for cross-verification of their accuracy.

Third, in all these districts the teachers and administrators collaborated to develop a common understanding of the processes to be used in the evaluation and the

About ERIC

The Educational Resources Information Center (ERIC) is a nationwide network operated by the Office of Educational Research and Improvement, U.S. Department of Education. ERIC serves educators by disseminating research results and other resource information that can be used in developing more effective educational programs.

The ERIC Clearinghouse on Educational Management, one of several such units in the system, was established at the University of Oregon in 1966. The Clearinghouse and its components process research reports and journal articles for announcement in ERIC's index and abstract bulletins.

Besides processing documents and journal articles, the Clearinghouse prepares bibliographies, literature reviews, monographs, and other interpretive research studies on topics in its educational area.

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract no. 400-83-0013.

Prior to publication, the manuscript was submitted to the National Association of Elementary School Principals for critical review and determination of professional competence. The publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of NAESP, the Clearinghouse, or the Department of Education.

ERIC Clearinghouse on Educational Management, University of Oregon, 1787 Agate Street, Eugene, Oregon 97403

goals involved. Fourth, the evaluation process and support systems were compatible with each other and with the district's overall goals and organizational context.

Perhaps the most significant finding to emerge from this study is the importance of teacher participation in developing and carrying out the evaluation process. The authors recommend that school districts involve expert teachers in peer support and assistance, that they involve teacher organizations in the design and oversight phases, and that they adopt a professional (as opposed to bureaucratic) approach to evaluation. Teachers should be held accountable not to arbitrary and hierarchically enforced criteria, they note, but rather to constantly evolving standards of practice developed by consensus among the teachers themselves.

3 McGreal, Thomas L. *Successful Teacher Evaluation.* Alexandria, Virginia: Association for Supervision and Curriculum Development, 1983. 175 pages. ED 236 776.

Thomas McGreal has worked with some 360 school districts and 75,000 supervisors and teachers around the country to build local teacher evaluation systems. In keeping with his conviction that teacher evaluation should be tailored to the specific needs of local districts, McGreal does not attempt to "sell" one approach or another. Rather, he organizes his book around eight "commonalities" of effective practices.

The first three commonalities constitute the framework for building a teacher evaluation system: an appropriate attitude, an evaluation model attuned to the desired purpose, and the separation of administrative and supervisory behavior. Evaluation, McGreal insists, should serve primarily to help teachers become more effective rather than to obtain documentation for personnel decisions. Systems that emphasize accountability tend to promote hostility between teachers and administrators, he says, whereas systems geared to improving instruction will cultivate mutual trust while still providing adequate information for assessing minimum competence when such assessment becomes necessary.

The next four commonalities are grouped as "focusing activities." The first is goal-setting as a cooperative activity between teacher and supervisor. Evaluation criteria that result from cooperative goal setting are more likely to fit the teacher's needs and methods, McGreal notes, whereas the alternative—imposition of criteria by the supervisor—forces all teachers to conform to an artificial standard. Cooperation also fosters mutual trust in the supervisory relationship.

Second, McGreal emphasizes the need for a common framework and a similar set of definitions about teaching from which to work. Third, to accommodate time constraints, he suggests two ways to increase the reliability of classroom observations: (1) narrowing the range of things to look for in accordance with previously established goals, and (2) increasing the amount of information assembled by the evaluator prior to the observation. McGreal also advocates the use of alternative sources of data to supplement classroom observations. These can include self-evaluation, peer evaluation, parent evaluation, student evaluation, student performance, and collection of such instructional artifacts as study guides, homework assignments, and tests. All these approaches, he notes, require a close working relationship between teachers and supervisors.

The final section focuses on training the staff and starting the system. As McGreal points out, the effectiveness of any evaluation system depends on the amount of training received by the participants—including both teachers and supervisors—in the skills and knowledge necessary to implement the system and effectively maintain it. An appropriate training program, he says, should include goal-setting skills for both supervisors and teachers, theoretical and practical training in the selected teaching focus, explanation and practice in the use of student descriptive data and artifact collection, and classroom observation and conference skills for supervisors.

4 Raths, James, and Hallie Preskill. "Research Synthesis on Summative Evaluation of Teaching." *Educational Leadership* 39, 4 (January 1982), 310-13. EJ 257 910.

Most educators agree that the primary aim of teacher evaluation should be to improve instruction. Nevertheless, school districts still need a reliable basis on which to make decisions on tenure, promotion, reassignment, or dismissal. For this reason, teacher evaluation systems need not only to help teachers improve, but also provide a basis for judging their performance. Raths and Preskill provide some clear thinking on this summative (as opposed to formative) dimension of teacher evaluation.

Judgment of teaching performance is inescapably a subjective process, the authors believe, since objective or measurable standards do not exist for the various aspects of teaching performance. Accordingly, some evaluators tend to emphasize tangible patterns, such as instructional techniques, teacher behavior, and lesson planning. Others emphasize such intangible qualities as cooperation, professionalism, and other general character traits. Whatever standards they stress, the authors say, evaluators must be able to communicate clearly the rationale behind their judgments to teachers, school boards, or courts.

Raths and Preskill describe two general approaches to summative evaluation. The "arithmetic" approach involves assigning quantitative values to relevant dimensions of instruction, weighting them for importance, rating teachers on each dimension, and then tallying the score. This procedure is popular, they note, because it confers the appearance of objectivity on an essentially subjective process.

The holistic approach involves summary assessments made by one or more evaluators and based on all available information. Despite the seeming objectivity of quantitative values, the authors say studies have shown that holistic judgments about teaching have more predictive validity than do arithmetic approaches.

Three methods are described for making holistic judgments: (1) a balloting system based on paired comparisons in which teachers are ranked from most effective to least; (2) a relative rating of teachers on a 1 to 5 basis, with 3 denoting the norm; and (3) the application of a similar five-point relative scale to a breakdown of separate teaching skills and behaviors.

The authors conclude that the inevitably subjective and impressionistic nature of summative evaluation can best be controlled by "triangulation" procedures, whereby judgments are made by several evaluators—including both supervisors and peers—and then cross-referenced.

5 Peterson, Donovan. "Legal and Ethical Issues of Teacher Evaluation: A Research-Based Approach." *Educational Research Quarterly* 7, 4 (Winter 1983), 6-16. EJ 284 820.

Use of teacher evaluation for personnel decisions necessarily has legal ramifications. In numerous court cases, teachers have challenged dismissals on grounds that administrators used unsound criteria or procedures to obtain evidence of incompetence.

Peterson's guide covers the legal issues that pertain to the design and plausibility of teacher evaluation systems. He offers helpful guidelines on such issues as due process, discrimination validity, reliability, high and low inference variables, representative observation of teacher behavior, research-based variables as opposed to those that are consensus-based, and number and length of observations.

The test of validity, for example, assesses whether teacher evaluation criteria are job-related and whether evaluation systems actually measure the attributes they claim to measure. Reliability refers to the consistency of judgments over time among separate observers. A variable is rated as being of high or low inference, depending on the degree of personal judgment an observer must apply to determine the presence or absence of the attribute for which the teacher is being rated, and its quality. High inference variables tend to lack reliability, Peterson says, whereas low inference variables—specific and measurable behaviors—tend to lack validity in distinguishing between effective and ineffective teaching.

Peterson suggests that schools adopt a three-level evaluation system. The first level, applying to the majority of experienced teachers, would consist of formative evaluation alone—for example, peer and self-evaluations, goal-setting, and an annual conference with a supervisor.

The second level would be for new teachers and those identified as perhaps needing improvement. The emphasis is still on the formative but involves several observations by trained observers to identify problems and develop remediation strategies.

The third level, summative evaluation, would be used only on teachers judged to be incompetent. Here documentation becomes important, along with judgments by more than one person and ample provision for remediation and assistance. By restricting summative evaluation to "borderline" cases, schools simultaneously reduce the potential for legal challenges and concentrate their resources for adequate documentation in the event of such challenges.

Research Roundup is issued as a service to members. Copies may be purchased from the National Association of Elementary School Principals, 1615 Duke Street, Alexandria, Virginia 22314. One copy, $2.00; 2-9 copies $1.50 each; 10 or more copies $1.00 each. Virginia residents should add 4% sales tax. **Payment must accompany order.**

R·E·S·E·A·R·C·H roundup

Teacher Evaluation

1615 Duke Street
Alexandria, Virginia 22314

(ISSN 8755-2590)

Nonprofit Org
U.S. Postage
PAID
Washington, D.C.
Permit #8039